Fusion of Probabilistic Evidence

Wei-min Dong and Wei-ling Chiang
Stanford University
Felix S. Wong
Weidlinger Associates, Los Altos

ABSTRACT

Fusion refers to the combination of two or more probability assignments to pieces of evidence that support the same hypotheses. The probability assignments usually result from different inference paths in reasoning and are, in general, different. Given a set of probability assignments for evidence to be fused, it is well known that certain constraints, called consistent bounds, must be satisfied. These bounds arise from the theory of probability and define an admissible domain for the fused evidence. However, because the bounds are, in general, interactive, a general methodology for computing the admissible domain other than a brute-force numerical approach (linear programming) is lacking. This paper examines the role of interaction in evidence fusion and demonstrates the effect of interaction on the fused evidence. A simple case consisting of one hypothesis supported by two pieces of evidence is considered, and the interactive bounds and admissible domain are derived analytically. In particular, the effect of different dependency assumptions on the consistent bounds is derived to show that the assumption of conditional dependence can lead to inconsistencies under certain circumstances, that is, the fused evidence lies outside the admissible domain. Uncertain evidence expressed in the form of bounds is not very useful in practice because the bounds tend to become large as the uncertainty is cascaded from one level to another in inferencing. Point evidence may be more helpful. Three suggestions for obtaining consistent point estimates from the consistent bounds are presented, and numerical examples are given.

KEYWORDS: ***evidence fusion, uncertainty combination, probability***

INTRODUCTION

With the rapid development and application of artificial intelligence techniques, there is a growing recognition of the need for uncertainty management.

Address correspondence to Felix Wong, Weidlinger Associates, 4410 El Camino Real, Suite 110, Los Altos, California 94022.

International Journal of Approximate Reasoning 1989; 3:437–466

655 Avenue of the Americas, New York, NY 10010 0888-613X/89/$3.50

On the commercial side, many expert system shells have been made available, for example, NEXPERT, PC Easy, EXSYS, and GURU, but they each favor a different strategy for treating uncertainties. However, most of the strategies used in commercial shells are ad hoc; they do not have a strong theoretical basis.

On the research side, many approaches to reasoning under uncertainty have been proposed and pursued. Notable examples include confirmation theory in MYCIN (Shortliffe [1]), Bayesian probability in PROSPECTOR (Duda et al. [2]), and evidence theory proposed by Dempster [3] and Shafer [4]. MYCIN uses confirmation theory to combine uncertainties in evidence to obtain the certainty factor for the hypothesis. As pointed out by Heckerman [5], the original definition of certainty factor is inconsistent with the combination functions used in MYCIN, and the assumptions implicit in the model are rarely true in practical applications. PROSPECTOR applies the Bayesian approach to propagate uncertainties in an inference network based on the assumption that all pieces of evidence are conditionally independent. As the most recent publications can attest, the Dempster-Shafer theory of evidence is getting more attention. However, it has been questioned by many researchers for the justification of its combination rule and its extension to rule-based inference (Zadeh [6, 7], Hunter [8], Yen [9]).

Hence, many theoretical and practical problems in uncertainty management remain to be solved. The main controversy appears to be evidence fusion—how evidence from different sources but supporting the same hypothesis should be combined.

This paper attempts to delineate evidence fusion from the point of view of probability theory. Fusion refers to the combination of two or more probability assignments to pieces of evidence that support the same hypotheses. The probability assignments usually result from different inference paths in reasoning and are, in general, different. Given a set of probability assignment for evidence to be fused, it is well known that certain constraints, called *consistent bounds,* must be satisfied. The consistent bounds arise from the theory of probability and define an admissible domain for the fused evidence. However, because the bounds are, in general, interactive—that is, the admissible value of one component depends on that of another—a general methodology for computing the admissible domain other than a brute-force numerical approach (linear programming) is lacking.

To delineate the interactive bounds and the way in which interaction affects the fused evidence, we consider a simple case consisting of one hypothesis supported by two pieces of evidence. For this simple case, the interactive bounds and admissible domain can be derived analytically. In particular, the effect of different dependency assumptions on the consistent bounds is derived to show that the assumption of conditional dependence can lead to inconsistencies under certain circumstances; that is, the fused evidence lies outside the admissible domain.

The paper is organized as follows. We summarize briefly several of the previous works on the subject and describe their relationship to our work presented in this paper. The issue of interaction and why it is important is then presented. To make the discussion more concrete, we analyze the admissible domain in probability space for the simple case of one hypothesis and two pieces of evidence. Consistent (but interactive) bounds on evidence fusion are derived. Next, different assumptions on probabilistic dependency are investigated to examine their effects on the admissible domain and the consistent bounds. Within this framework, the validity of the conditional independence assumption often quoted in the literature can be readily evaluated. Finally, methods for point estimation are suggested to make evidence fusion more computationally appealing. One method uses the midpoint of the consistent bounds as a point estimate, and another uses the centroid of the admissible domain. A third method is based on applying the principle of maximum entropy to the admissible domain.

PREVIOUS WORKS

Among the earliest applications of probability to evidence fusion, the most notable appears to be PROSPECTOR (Duda et al. [2]). The Bayesian approach is used to propagate uncertainties in an inference network based on the assumption that all pieces of evidence are conditionally independent. The independence assumption is used to simplify the computation in PROSPECTOR, but a side effect is that it has spurred many, more general studies on probabilistic fusion. In the following we review several that are most relevant to our work.

If one does not know (or impose an assumption on) the dependency of the evidence, the relationship can range from total dependency to mutual exclusivity. Consequently, it is well known from probability theory that a point estimate of the fused evidence is not possible. Rather, the possible dependence relations appear as constraints that restrict the fused evidence to be within an admissible domain. Probabilistic uncertainties are propagated as bounds that in turn lead to a linear programming problem. Many researchers have addressed this issue (see, for example, Good [10] and Cooper [11]).

However, one difficulty in computation that has not been addressed is that the bounds are interactive and the function is nonlinear. To be more specific, if E_1 and E_2 are two pieces of evidence supporting a hypothesis H and $\wedge$ denotes the intersection, Bayes's theorem gives

$$P(H \mid E_1 \wedge E_2) = \frac{P(H \wedge E_1 \wedge E_2)}{P(E_1 \wedge E_2)}$$

where the numerator and denominator are interactive intervals, that is, the admissible value of $P(H \wedge E_1 \wedge E_2)$ depends on the value of $P(E_1 \wedge E_2)$ and vice versa. The determination of $P(H|E_1 \wedge E_2)$ is a nonlinear programming problem and is addressed in this paper.

Prompted mostly by the work of Duda, Hart, and Nilsson in PROSPECTOR [2], much research has gone into investigating the ramifications of making the conditional independence assumption and the consistency (or inconsistency) that results according to probability theory. The reason for such interest is this: In expert system inference where evidence is expressed in terms of subjective probabilities, evidence may or may not be consistent and Bayesian updating may or may not be feasible, depending on the number of hypotheses addressed and whether these hypotheses are mutually exclusive or exhaustive. For example, Pednault et al. [12] set out to show that the assumption of independence of the evidence under both a hypothesis and its negation leads to inconsistency when there are three or more mutually exclusive and exhaustive hypotheses. Subsequently, Glymour [13] presented a counterexample to disprove Pednault's claim. Johnson [14] reexamined the problem and showed that when there are three or more mutually exclusive and jointly exhaustive hypotheses and independence under both hypothesis and negation is assumed, then at most one evidence event can affect the probability of any hypothesis. For mutually exclusive and jointly exhaustive hypotheses, independence under hypothesis (but not, simultaneously, independence under the hypothesis's negation) is acceptable. That is, if $E_1, \ldots, E_n$ are the evidence events, then for each hypothesis H_i we may assume

$$P(E_1, \ldots, E_n|H_i) = \prod_j P(E_j|H_i)$$

However, we get into trouble by making the additional assumption that

$$P(E_1, \ldots, E_n|\neg H_i) = \prod_j P(E_j|\neg H_i)$$

as well, where $\neg H_i$ denotes the negation of H_i

In a classical Bayesian problem, one works with a number of mutually exclusive and jointly exhaustive hypotheses, a set of priors over the hypotheses, and the conditional probabilities $P(E_j|H_i)$. The conditional independence assumption $P(E_1, \ldots, E_n|H_i)$ is then sufficient for updating and solution of the problem. Adding the independence assumption on the negation, $P(E_1, \ldots, E_n|-{}_1H_i)$, imposes another condition that makes the problem overdetermined. However, in expert system inference involving subjective probabilities, it is not always possible to work with hypotheses that are mutually exclusive and jointly exhaustive (e.g., medical diagnostics). It is of interest to know how many more conditions (dependency assumptions) are required for the solution of the problem in that case. Specifying too many conditions leads to an overdetermined situation as discussed by Pednault, Glymour, and Johnson. Specifying too few leads to an underdetermined situation described by bounds and nonlinear programming as discussed by Good and Cooper.

In [15] Nilsson establishes a beautiful framework for probabilistic logic that is based on a semantical generalization of logic in which the truth values are probability values between 0 and 1. In particular, he shows how the framework can be used to compute probabilities conditioned on additional information, namely, Bayesian updating. He mentions in passing that when the probabilities are given as bounds the conditional probabilities will also be bounded and that the difference of the upper and lower bounds expresses our ignorance. However, no details on the computation of the bounds on the conditional probability are given, and the difficulty due to interaction of the bounds is not addressed.

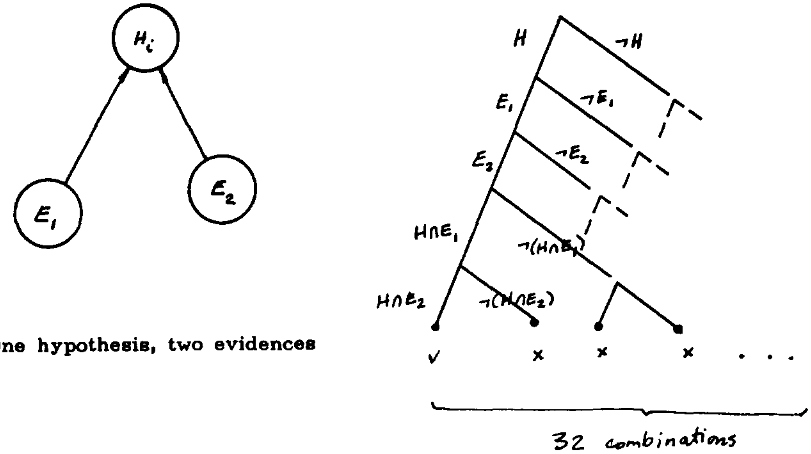
In the section we shall make use of Nilsson's framework to delineate the effect of dependence assumptions on probability fusion. We shall work with the simple case of one hypothesis and two pieces of evidence and show that the issue of consistent evidence assignments can be reduced to a comparison of the number of unknowns versus the number of conditions in Nilsson's possible world. For the underdetermined situation, we further delineate the interaction among the probability bounds and the admissible domain for the conditioned probability. In the general case, Nilsson's world gives the number of additional conditions that must be imposed (or can be imposed) to make the problem determinate. However, unlike the simple case considered here, an analytic solution for the general case is not yet possible.

POSSIBLE WORLD AND PROBABILISTIC LOGIC

We use the example of one hypothesis H and two pieces of evidence E_1 and E_2 to introduce Nilsson's possible world. As a first step, we establish the binary tree shown in Figure 1, starting with the two possible logical values for H: H is true (H) and H is false ($\neg H$). Each of these two branches leads to two more branches, E_1 is true and E_1 is false, and so forth. Since there are five variables of interest (H, E_1, E_2, $H \wedge E_1$, $H \wedge E_2$), which we shall denote collectively by the vector $\mathbf{X}$, and each variable has two states, there are $2^5 = 32$ possible combinations of $\mathbf{X}$ that correspond, respectively, to the 32 leaf nodes of the binary tree. Obviously, not all leaf nodes can be reached, because certain combinations are not logically consistent [e.g., H true and ($H \wedge E_1$) false]. We mark the logical ones with a check mark ($\surd$) and the illogical ones with a cross ($\times$) in Figure 1.

The binary logic tree can be summarized more succinctly as a matrix, such as the matrix $\mathbf{V}$, when 1 is used to denote the true state and 0 the false state.

$$\mathbf{V} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \leftrightarrow \begin{array}{l} H \\ E_1 \\ E_2 \\ H \wedge E_1 \\ H \wedge E_2 \end{array} \tag{1}$$



(a) One hypothesis, two evidences

(b) Binary logic tree

Figure 1. Possibility logic and binary tree for one hypothesis and two pieces of evidence.

Each column of $\mathbf{V}$ corresponds to one possible value of $\mathbf{X}$ (or a leaf node of the tree), and it is clear that there are only eight logically consistent vectors for the example. Consequently, only the logically consistent combinations are shown in Eq. (1), and they constitute the possible world. The eight combinations are also illustrated in the Venn diagram of Figure 2. The possible world consists of eight states denoted by $p_1, \ldots, p_8$ in the figure; the state p_1 corresponds to the first column of $\mathbf{V}$, p_2 the second column, and so on.

The general problem to be solved is as follows. Given the probability assignments $p(H)$, $p(E_1)$, $p(E_2)$, $p(H \wedge E_1)$, and $p(H \wedge E_2)$, which we shall group together as a vector and call $\boldsymbol{\pi}$, find the probabilities associated with the eight logically admissible states $p_1, \ldots, p_8$. Without causing confusion, we have denoted the probabilities by the names of the states (i.e., $p_1, \ldots, p_8$, respectively) as well. We shall also group them together in a vector called $\mathbf{p}$. It is understood that

$$\mathbf{p} \geqslant \mathbf{0} \tag{2}$$

because probabilities are by definition non-negative.

The unknown vector $\mathbf{p}$ and the known vector $\boldsymbol{\pi}$ are then related by the matrix $\mathbf{V}$ as

$$\boldsymbol{\pi} = \mathbf{V}\mathbf{p} \tag{3}$$

which is a linear system of five equations with eight unknowns. Since the probability assignments must sum to 1 according to probability theory, an

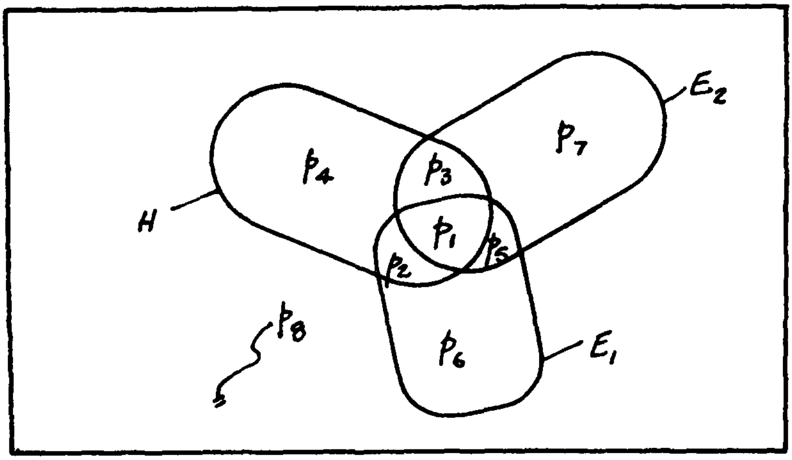


Figure 2. Possible world for one hypothesis and two pieces of evidence.

additional equation in the form of

$$\sum_i p_i = 1 \tag{4}$$

makes a total of six equations. Hence, the system of Eqs. (2)–(4) is underdetermined and has two degrees of freedom. A point solution cannot be found unless two more conditions on $\mathbf{p}$ are added. As such, these conditions constitute the constraints on the solution set.

The foregoing discussion applies equally well to more complex situations involving more than one hypothesis and three or more pieces of evidence. *The resultant set of equations is always underdetermined, and the equations serve as constraints on the solution set.* When the required number of additional equations are prescribed, usually through dependence assumptions (i.e., on the E_i's) and conditional dependence assumptions (i.e., on the E_i's conditioned on H being true or false), point solutions will then be possible. For example, the two degrees of freedom in Eqs. (2)–(4) can be removed by assuming that E_1 and E_2 are unconditionally independent and independent conditioned on H.

Because our ultimate goal is to find $p(H|E_1 \wedge E_2)$, which is related to $p(E_1 \wedge E_2)$ and $p(H \wedge E_2 \wedge E_2)$, we consider two subworlds of the possible world. The same probabilistic logic framework applies to these subworlds.

SUBWORLDS 1: $p(E_1)$, $p(E_2) \rightarrow p(E_1 \wedge E_2)$

The possible world for the subset $p(E_1)$, $p(E_2) \rightarrow p(E_1 \wedge E_2)$ consists of four states as shown below and in the Venn diagram of Figure 3.

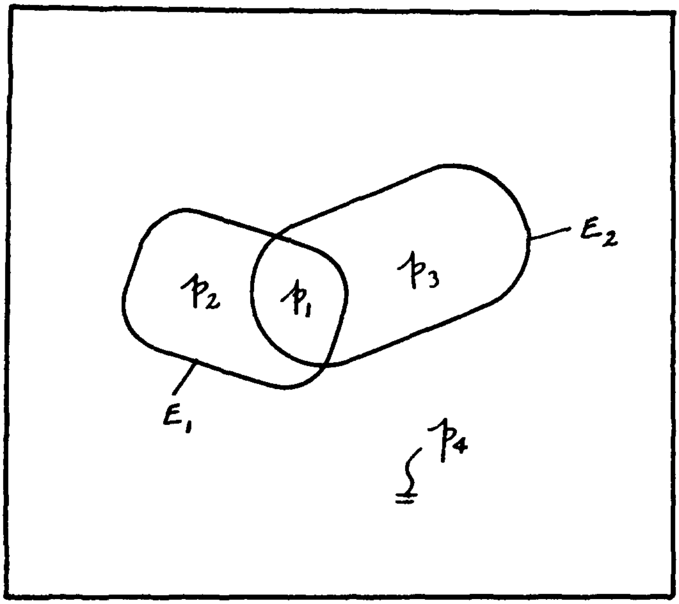


Figure 3. Possible world for subworld 1.

$$\begin{array}{c} E_1 \\ E_2 \end{array} \rightarrow \begin{bmatrix} \mathrm{T} & \mathrm{T} & \mathrm{F} & \mathrm{F} \\ \mathrm{T} & \mathrm{F} & \mathrm{T} & \mathrm{F} \end{bmatrix}$$

Hence,

$$\mathbf{V} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \tag{5}$$

The given are

$$\boldsymbol{\pi} = \{p(E_1), p(E_2)\}^T$$

which will be represented in shorthand as

$$\boldsymbol{\pi} = \{x, y\}^T \tag{6}$$

We use the superscript T to denote a column vector. The unknown $\mathbf{p}$ vector is

$$\mathbf{p} = \{p_1, p_2, p_3, p_4\}^T \tag{7}$$

where p_1, p_2, p_3, and p_4 refer to the four parts in Figure 3. We are interested mainly in finding p_1, or $p(E_1 \wedge E_2)$, which we also denote as z for convenience.

It can be shown that Eqs. (2), (3), and (5)–(7) reduce to the following

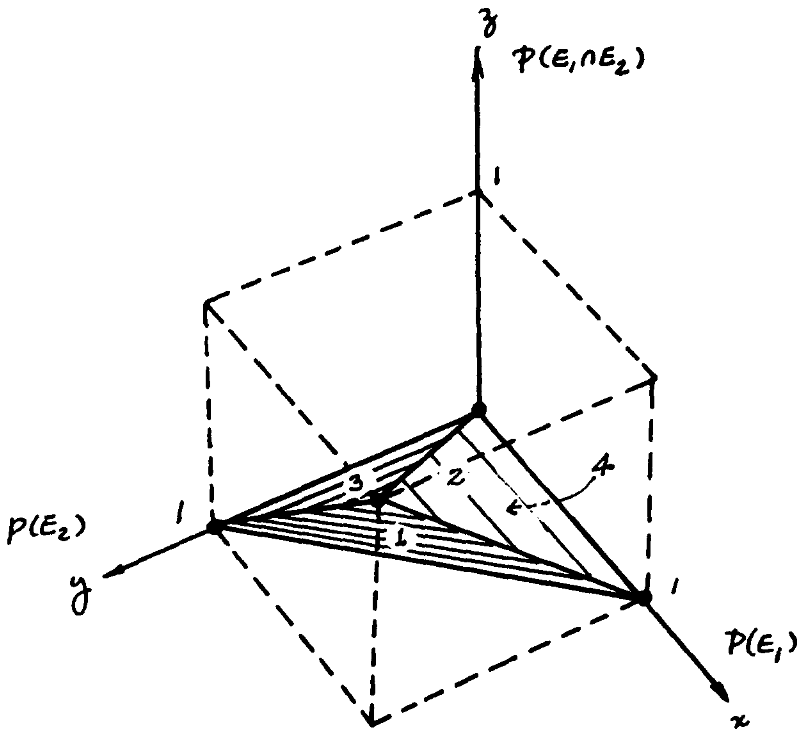


Figure 4. Convex hull for subworld 1.

inequalities:

$$z \geqslant 0$$
$$x \geqslant z$$
$$y \geqslant z$$
$$x+y-z \leqslant 1 \tag{8}$$

which define an admissible domain in (x, y, z) space given in Figure 4. The admissible domain is a tetrahedron, bounded by the four planes

$$S_1:\ x+y-z=1$$
$$S_2:\ y-z=0$$
$$S_3:\ x-z=0$$
$$S_4:\ z=0 \tag{9}$$

as noted in Figure 4

Nilsson calls this admissible domain the convex hull of $\boldsymbol{\pi}$. Note that the hull is anchored by four points that correspond to the column vectors of $\mathbf{V}$, which are also the $\boldsymbol{\pi}$ vectors that correspond to the extreme vectors of $\mathbf{p}$, namely, $\mathbf{p} = \{1, 0, 0, 0\}^T$, $\{0, 1, 0, 0\}^T$, $\{0, 0, 1, 0\}^T$, and $\{0, 0, 0, 1\}^T$. It emphasizes the fact that consistent values for the probabilities of E_1, E_2, and $E_1 \wedge E_2$ must lie in the convex hull anchored by the extreme values of the vector $\{E_1, E_2, E_1 \wedge E_2\}^T$.

As a side interest, the constraints of Eq. (8), which are denoted by the convex hull of Figure 4, can be summarized succinctly as

$$\max(x+y-1, 0) \leqslant z \leqslant \min(x, y)$$

or as

$$\max[p(E_1)+p(E_2)-1, 0] \leqslant p(E_1 \wedge E_2) \leqslant \min[p(E_1), p(E_2)] \quad (10)$$

which are the well-known bounds on $p(E_1 \wedge E_2)$ given that the dependence relation between E_1 and E_2 is unknown. The lower bound corresponds to the minimal dependence condition between E_1 and E_2, and the upper bound corresponds to the maximal dependence condition.

SUBWORLD 2: $p(H)$, $p(H \wedge E_1)$, $p(H \wedge E_2) \rightarrow p(H \wedge E_1 \wedge E_2)$

This subworld has three given probabilities, namely, the probability of H, $H \wedge E_1$, and $H \wedge E_2$. The truth combinations of these states are listed below:

$$\begin{array}{l} H \\ H \wedge E_1 \rightarrow \\ H \wedge E_2 \end{array} \begin{bmatrix} \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{F} \\ \mathrm{T} & \mathrm{T} & \mathrm{F} & \mathrm{F} & \mathrm{F} \\ \mathrm{T} & \mathrm{F} & \mathrm{T} & \mathrm{F} & \mathrm{F} \end{bmatrix} \quad (11)$$

The possible world is given in Figure 5, which defines p_1 through p_5. Again, p_1 or $p(H \wedge E_1 \wedge E_2)$ is of interest.

The matrix equation is

$$\mathbf{V}\mathbf{p} = \boldsymbol{\pi} \quad (12)$$

or

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{Bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{Bmatrix} = \begin{Bmatrix} p(H) \\ p(H \wedge E_1) \\ p(H \wedge E_2) \end{Bmatrix} \quad (13)$$

$$= p(H) \begin{Bmatrix} 1 \\ p(H \wedge E_1)/p(H) \\ p(H \wedge E_2)/p(H) \end{Bmatrix} \quad (14)$$

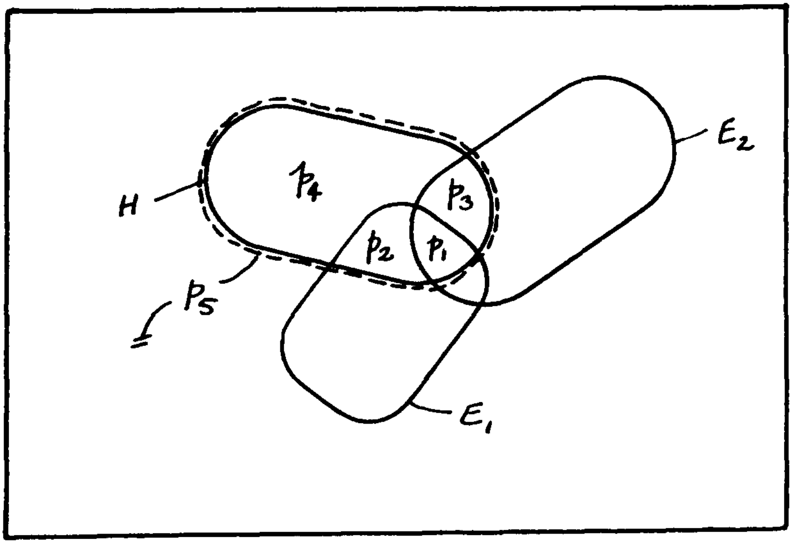


Figure 5. Possible world for subworld 2.

We can plot the convex hull in the scaled three-dimensional space $[p(H \wedge E_1)/p(H), p(H \wedge E_2)/p(H), p(H \wedge E_1 \wedge E_2)/p(H)]$ directly if we note that V can be decomposed into

$$\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{array}$$

where the submatrix at the lower left is identical to the one in Eq. (5) and the one at the upper left is simply Eq. (2). Hence, the convex hull for this subworld, when dimensioned by $p(H)$, is identical to the one considered previously and is shown in Figure 6. The admissible domain can be summarized succinctly by

$$\max[x+y-p(H), 0] \leqslant z \leqslant \min(x, y)$$

or

$$\begin{aligned} &\max[p(H \wedge E_1)+p(H \wedge E_2)-p(H), 0) \\ &\leqslant p(H \wedge E_1 \wedge E_2) \\ &\leqslant \min[p(H \wedge E_1), p(H \wedge E_2)] \end{aligned} \tag{15}$$

which is also a well-known result in probability.

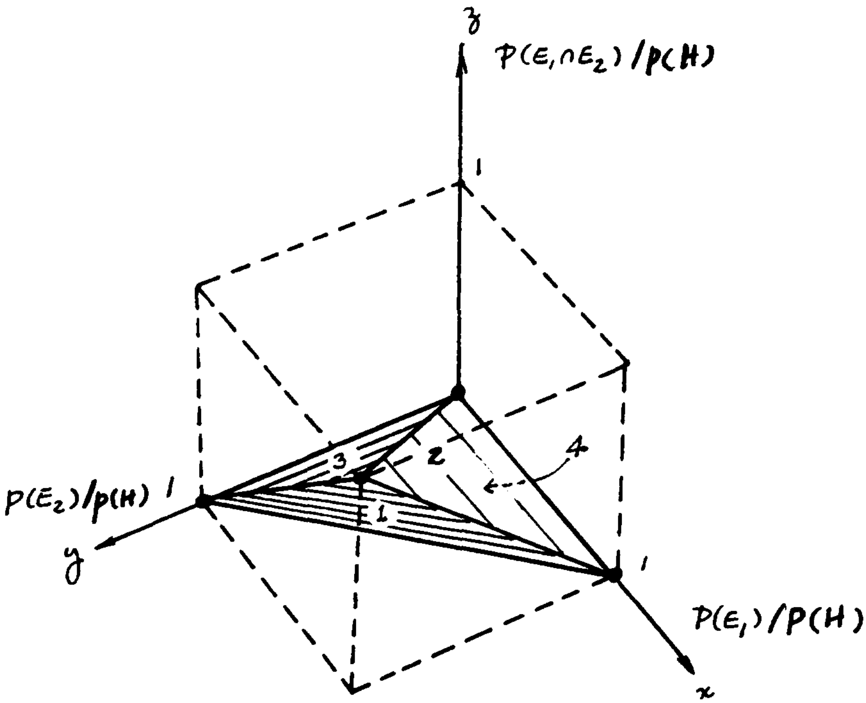


Figure 6. Convex hull for subworld 2.

INTERACTIVITY BETWEEN THE TWO SUBWORLDS

According to Bayes's theorem,

$$p(H|E_1 \wedge E_2) = \frac{p(H \wedge E_1 \wedge E_2)}{p(E_1 \wedge E_2)} \tag{16}$$

Hence, $p(H|E_1 \wedge E_2)$ is known when $p(E_1 \wedge E_2)$ and $p(H \wedge E_1 \wedge E_2)$ are known. We have shown previously that $p(E_1 \wedge E_2)$ and $p(H \wedge E_1 \wedge E_2)$ can be computed only to within certain bounds from $p(H)$, $p(H \wedge E_1)$, $p(H \wedge E_2)$, $p(E_1)$, and $p(E_2)$ [by Eqs. (10) and (15), respectively], since the dependency between E_1 and E_2 and that among E_1, E_2, and H are now known. It follows that $p(H|E_1 \wedge E_2)$, can also be computed by Eq. (16) only to within certain bounds.

If $p(E_1 \wedge E_2)$ and $p(H \wedge E_1 \wedge E_2)$ can vary freely within their respective bounds, the computation according to Eq. (16) is trivial. However, $p(E_1 \wedge E_2)$ *and* $p(H \wedge E_1 \wedge E_2)$ *are not independent*, because

$$p(E_1 \wedge E_2) = p(H \wedge E_1 \wedge E_2) + p(\neg H \wedge E_1 \wedge E_2) \tag{17}$$

Consequently, the possible range of $p(E_1 \wedge E_2)$ as given by Eq. (10) depends on the value of $p(H \wedge E_1 \wedge E_2)$ and vice versa. We say that $p(E_1 \wedge E_2)$ and $p(H \wedge E_1 \wedge E_2)$ are *interactive* in view of Eq. (17).

Interaction exists among subworlds of Nilsson's possible world approach and has not been addressed explicitly. This issue was also ignored in previous studies on the subject of evidence fusion. However, it is a crucial step because Bayes's equation [Eq. (16)] always involves interactive subworlds, and unless a satisfactory resolution of this difficulty can be found, much of the advances in fusion research cannot be fully utilized. In a later section we show how interaction affects the admissible domain and the bounds on the fused evidence.

INTERACTIVE ADMISSIBLE DOMAIN

For the example being considered, the constraint given by Eq. (17) can be rewritten as

$$\begin{aligned} & p(H \wedge E_1 \wedge E_2)+p(\neg H \wedge E_1 \wedge E_2)_{\min} \\ & \quad \leqslant p(E_1 \wedge E_2) \\ & \quad \leqslant p(H \wedge E_1 \wedge E_2)+p(\neg H \wedge E_1 \wedge E_2 \wedge \rangle_{\max} \end{aligned} \tag{18}$$

where

$$p(\neg H \wedge E_1 \wedge E_2)_{\min}=\max[0, p(E_1 \wedge \neg H)+p(E_2 \wedge \neg H)-p(\neg H)]$$

and

$$p(\neg H \wedge E_1 \wedge E_2)_{\max}=\min[p(E_1 \wedge \neg H), p(E_2 \wedge \neg H)]$$

Hence, Eq. (18) becomes

$$\begin{aligned} & p(H \wedge E_1 \wedge E_2)+\max[0, p(E_1 \wedge \neg H)+p(E_2 \wedge \neg H)-p(\neg H)] \\ & \leqslant p(E_1 \wedge E_2) \\ & \leqslant p(H \wedge E_1 \wedge E_2)+\min[p(E_1 \wedge \neg H), p(E_2 \wedge \neg H)] \end{aligned} \tag{19}$$

Because

$$p(E_1 \wedge \neg H)=p(E_1)-p(E_1 \wedge H)$$

$$p(E_2 \wedge \neg H)=p(E_2)-p(E_2 \wedge H)$$

and

$$p(\neg H)=1-p(H)$$

Eq. (19) becomes

$$
\begin{aligned}
&p(H \wedge E_1 \wedge E_2) + \max[0, p(E_1) + p(E_2) + p(H) \\
&\qquad - p(E_1 \wedge H) - p(E_2 \wedge H) - 1] \\
&\quad \leqslant p(E_1 \wedge E_2) \\
&\quad \leqslant p(H \wedge E_1 \wedge E_2) + \min[p(E_1) - p(E_1 \wedge H), p(E_2) - p(E_2 \wedge H)] \qquad (20)
\end{aligned}
$$

Hence, when interaction is taken into consideration, Eq. (20) supersedes Eq. (10) and together with Eq. (16) defines a two-dimensional region in the $\{p(H \wedge E_1 \wedge E_2), p(E_1 \wedge E_2)\}$ plane, which we shall call the interactive admissible domain. Any point inside this domain leads to a possible value for $p(H|E_1 \wedge E_2)$ according to Eq. (16). Expressed in another way, the uncertainty on $p(H|E_1 \wedge E_2)$, the fused evidence, is a direct result of the uncertainty depicted by the interactive admissible domain.

NUMERICAL EXAMPLES

Several numerical examples of the interactive admissible domain are given in Figures 7–10, with each figure (domain) corresponding to a consistent set of probability assignments $p(H), \ldots, p(E_2)$ as indicated in the figure.[1]

The admissible domains are parallelograms, bounded by vertical lines corresponding to upper and lower bounds on $p(H \wedge E_1 \wedge E_2)$ and 45° lines. Note that when interaction is ignored, that is, Eq. (10) is used instead of Eq. (20), the "admissible" domains are rectangles formed by the two (independent) sets of bounds on $p(H \wedge E_1 \wedge E_2)$ and $p(E_1 \wedge E_2)$ as denoted by the dashed line in Figure 7 for comparison. The rectangle is larger than and contains the parallelogram. The effect of interaction is to make the admissible domain smaller with inclined boundaries.

[1] Note that if the assignments for the probabilities $p(H), \ldots, p(E_2)$ are subjective, as is usually the case in evidential reasoning, they should be checked for consistency. In particular

$$
\begin{aligned}
p(H) &= p(H|E_i)p(E_i) + p(H|\neg E_i)p(\neg E_i) \\
&= p(H|E_i)p(E_i) + p(H|E_i)[1 - p(E_i)], \qquad i = 1, 2
\end{aligned}
$$

$p(H|\neg E_i)$ is usually not given, but we know that it must be in the range [0, 1]. Hence,

$$
p(H|E_i)p(E_i) \leqslant p(H) \leqslant p(H|E_i)p(E_i) + [1 - p(E_i)], \qquad i = 1, 2
$$

When the prescribed set of prior $p(H)$ and posteriors $p(H|E_1)$, $p(H|E_2)$, and $p(E_1)$, $p(E_2)$ satisfies these inequalities, we say that the set is consistent. Naturally, different sets of given conditions are also possible, such as the set $p(H)$, $p(H|E_i)$, $p(H|\neg E_i)$, $p(\neg H|E_i)$, and $p(\neg H|\neg E_i)$, as long as they are consistent.

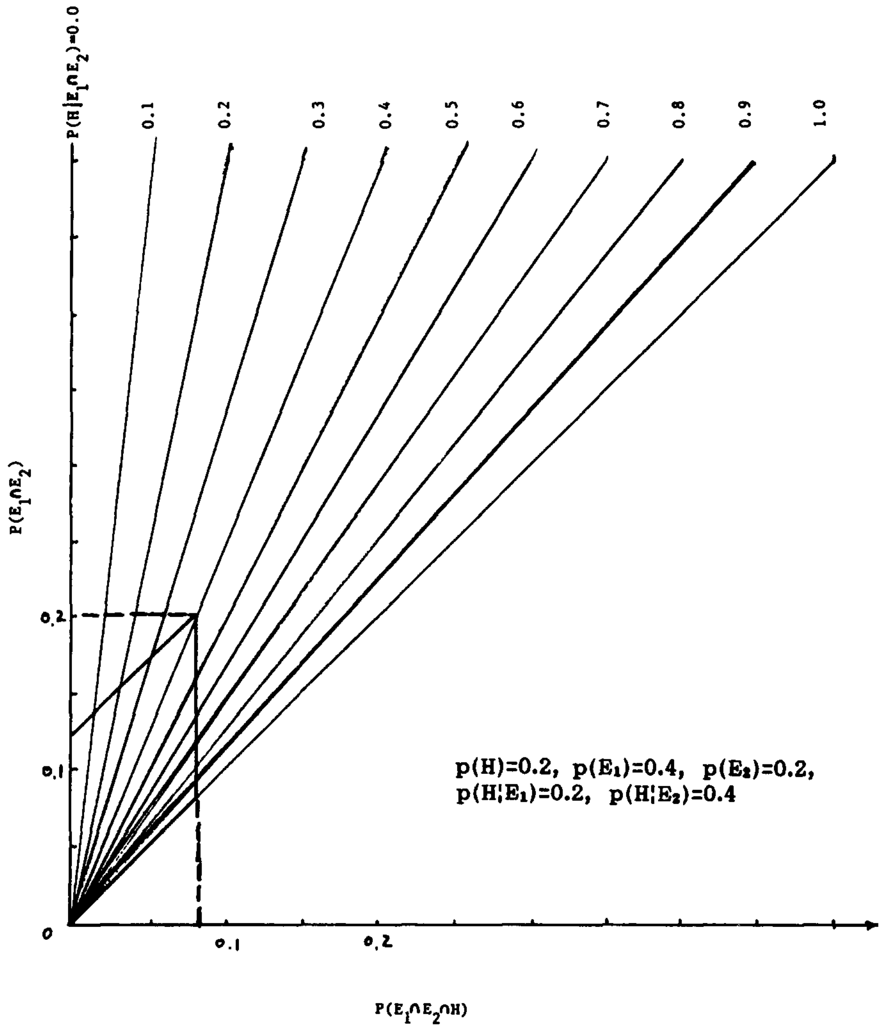


Figure 7. Admissible domain for Example 1.

Ignoring interaction has more deleterious effects than simply a wider bound. The bounds are often inadmissible and even undefined. We shall use the example in Figure 7 to illustrate. If interaction is ignored,

$$p(E_1 \wedge E_2) = [0, 0.2]$$

$$p(E_1 \wedge E_2 \wedge H) = [0, 0.08]$$

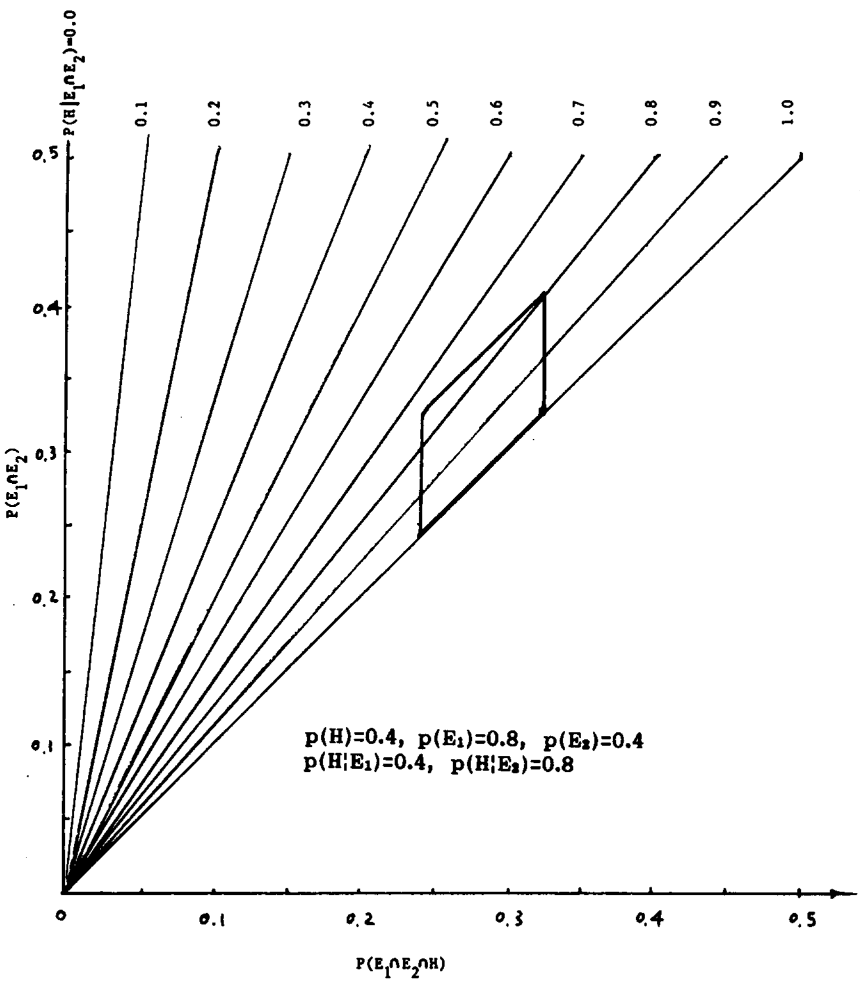


Figure 8. Admissible domain for Example 2.

Hence,

$$p(H|E_1 \wedge E_2) = \frac{p(E_1 \wedge E_2 \wedge H)}{p(E_1 \wedge E_2)} = \frac{[0, 0.08]}{[0, 0.2]}$$
$$= \text{``undefined''}$$

when considered as an interval division (e.g., when the numerator is 0.08 and the denominator is 0). On the other hand, the consistent bounds are [0, 1] as the figure shows.

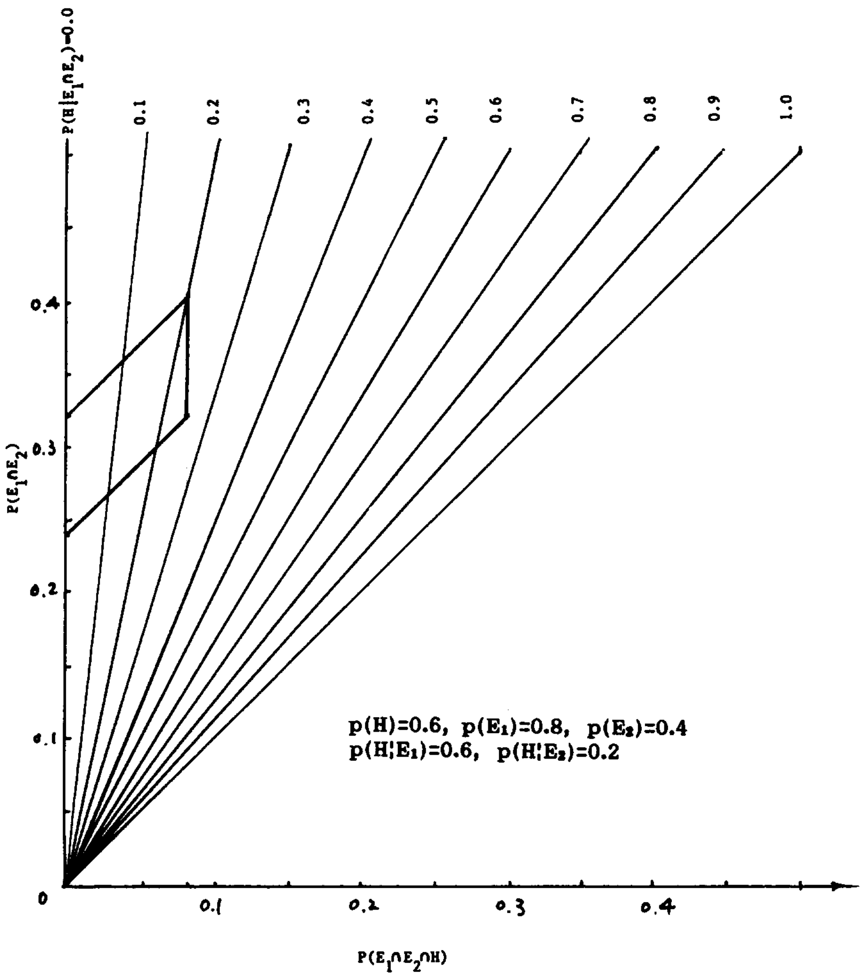


Figure 9. Admissible domain for Example 3.

As another example, refer to Figure 10.

$$p(E_1 \wedge E_2) = [0.4, 0.6]$$

$$p(E_1 \wedge E_2 \wedge H) = [0.12, 0.48]$$

and if interaction is ignored,

$$p(H | E_1 \wedge E_2) = \frac{[0.12, 0.48]}{[0.4, 0.6]} = [0.12, 1.2]$$

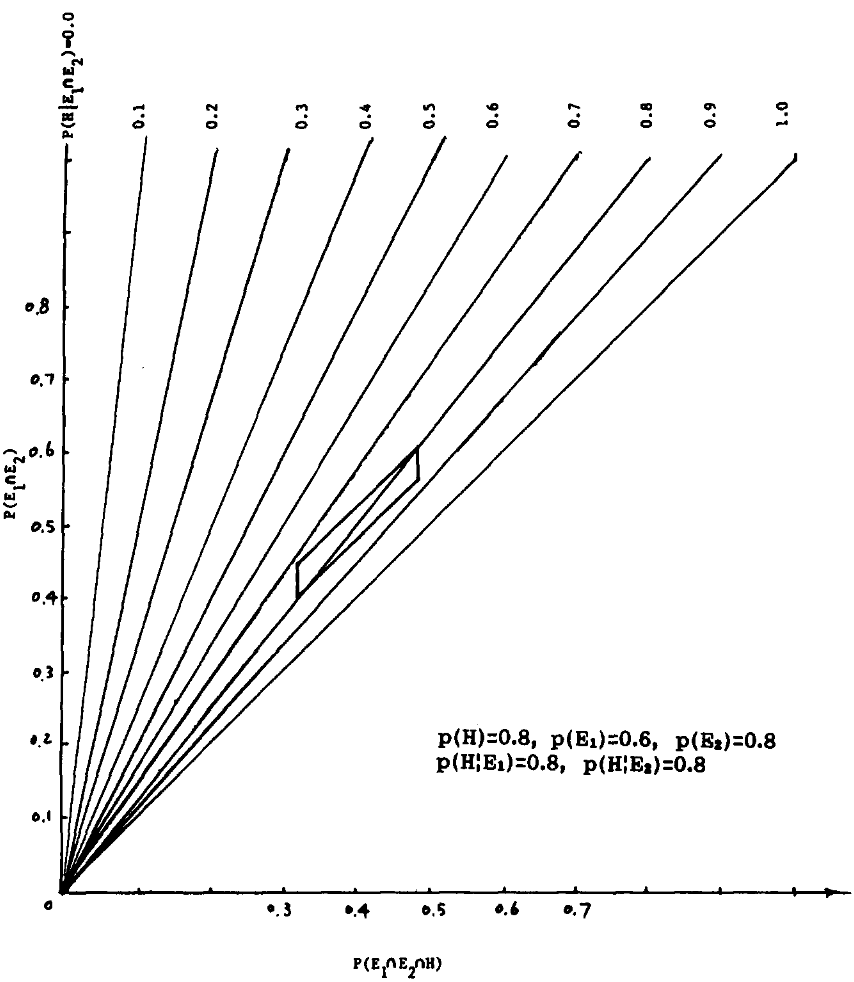


Figure 10. Admissible domain for Example 4.

which is an inadmissible result since probability must not exceed 1. The consistent bounds are [0.73, 0.86].

Intuitively, one can sense trouble coming in performing the simple interval divisions above. For any value of $p(E_1 \wedge E_2)$, say a, taken from the admissible interval of $p(E_1 \wedge E_2)$, we know that the corresponding value of $p(E_1 \wedge E_2 \wedge H)$, say b, which we should take from its admissible interval, should not be greater than a. Otherwise, the result of the interval division will exceed 1. Interaction formalizes that constraint.

To expedite subsequent development, we shall use the shorthand notation,

$$\begin{aligned}
\mathrm{HE12} &= p(E_1 \wedge E_2 \wedge H) \\
\mathrm{HE12}_{\min} &= \min[p(E_1 \wedge E_2 \wedge H)] \\
&= \max[p(E_1 \wedge H) + p(E_2 \wedge H) - p(H), 0], \qquad \text{from Eq. (15)} \\
\mathrm{HE12}_{\max} &= \max[p(E_1 \wedge E_2 \wedge H)] \\
&= \min[p(E_1 \wedge H), p(E_2 \wedge H)] \qquad \text{also from Eq. (15)} \\
\neg\mathrm{HE12} &= p(E_1 \wedge E_2 \wedge \neg H) \\
\neg\mathrm{HE12}_{\min} &= \min[p(E_1 \wedge E_2 \wedge \neg H)] \\
&= \max[p(E_1) + p(E_2) + p(H) - p(E_1 \wedge H) - p(E_2 \wedge H) - 1, 0] \\
\neg\mathrm{HE12}_{\max} &= \min[p(E_1 \wedge E_2 \wedge \neg H)] \\
&= \min[p(E_1) - p(E_1 \wedge H), p(E_2) - p(E_2 \wedge H)]
\end{aligned}$$

The admissible domain can then be defined by the bounds

$$\mathrm{HE12}_{\min} \leqslant p(E_1 \wedge E_2 \wedge H) \leqslant \mathrm{HE12}_{\max} \tag{21}$$

$$\mathrm{HE12} + \neg\mathrm{HE12}_{\min} \leqslant p(E_1 \wedge E_2) \leqslant \mathrm{HE12} + \neg\mathrm{HE12}_{\max} \tag{22}$$

CONSISTENT BOUNDS ON FUSED EVIDENCE

The previous section shows that given a set of consistent probability assignments for $p(H)$, $p(H|E_1)$, $p(H|E_2)$, $p(E_1)$, and $p(E_2)$, the values of $p(E_1 \wedge E_2 \wedge H)$ and $p(E_1 \wedge E_2)$ are restricted to the admissible domain. Hence, it is obvious that the lowest possible value for $p(H|E_1 \wedge E_2)$ as given by Eq. (16) is when $p(E_1 \wedge E_2 \wedge H)$ takes on its smallest value and $p(E_1 \wedge E_2)$ takes on its largest value within the admissible domain. That is the upper left vertex of the domain, or

$$p_L = \frac{\mathrm{HE12}_{\min}}{\mathrm{HE12}_{\min} + \neg\mathrm{HE12}_{\max}} \tag{23}$$

Similarly, the largest value for $p(H|E_1 \wedge E_2)$ is given by the lower right vertex of the admissible domain,

$$p_U = \frac{\mathrm{HE12}_{\max}}{\mathrm{HE12}_{\max} + \neg\mathrm{HE12}_{\min}} \tag{24}$$

The bounds p_L and p_u are called the consistent bounds for $p(H|E_1 \wedge E_2)$,[2] and

$$p_L \leqslant p(H|E_1 \wedge E_2) \leqslant p_U \tag{25}$$

The contours of constant $p(H|E_1 \wedge E_2)$ are the straight lines through the origin as shown in Figures 7–10.

To summarize the development described so far, the admissible domain and the consistent bounds for the fused evidence evolve because given any set of (consistent) probability assignments $p(H), \ldots, p(E_2)$, the joint probabilities $p(E_1 \wedge E_2)$ and $p(E_1 \wedge E_2 \wedge H)$ are not known exactly and the evidence cannot be combined with certainty. In particular, an interval value results for $p(E_1 \wedge E_2)$ and $p(E_1 \wedge E_2 \wedge H)$ due to ignorance of the relationship between E_1 and E_2 and between $E_1 \wedge H$ and $E_2 \wedge H$.

However, the interval constraints on $p(E_1 \wedge E_2)$ and $p(E_1 \wedge E_2 \wedge H)$ are not independent. Indeed, one of the major difficulties encountered in probabilistic evidence fusion that remains unresolved is the fact that the interval constraints are interactive. For the simple case considered here, the interaction is shown to lead to the definition of the admissible domain and bounds for $p(H|E_1 \wedge E_2)$ as given by Eq. (25). Any value of $p(H|E_1 \wedge E_2)$ within these bounds is possible and corresponds to certain compatible pairs of $p(E_1 \wedge E_2)$ and $p(E_1 \wedge E_2 \wedge H)$ in the admissible domain (and certain dependence conditions). By the same token, any value of $p(H|E_1 \wedge E_2)$ outside of these bounds will signify inconsistency somewhere in the assigned probabilities. Hence, the bounds p_L and p_U as given by Eqs. (23) and (24) are called consistent bounds for the fused evidence. In application, the consistent bounds may be used to safeguard consistency in the assigned probabilities and in the updating algorithms.

EFFECT OF DEPENDENCY ASSUMPTION ON CONSISTENT BOUNDS

We have seen from the possible world frameword that additional conditions (assumptions) will decrease the number of degrees of freedom in the computation of the fused evidence. In particular, when information regarding the dependency relation between E_1 and E_2 or between $E_1 \wedge H$ and $E_2 \wedge H$ is known, the uncertainty in fusion will be reduced. The result is a narrower set of consistent bounds for $p(H|E_1 \wedge E_2)$. In this section, the effect of different dependence assumptions on the admissible domain and the consistent bounds will be examined.

[2] We use the term *consistent bounds* here to denote that interaction between the subworlds have been taken into account in generating the admissible domain that leads to these bounds.

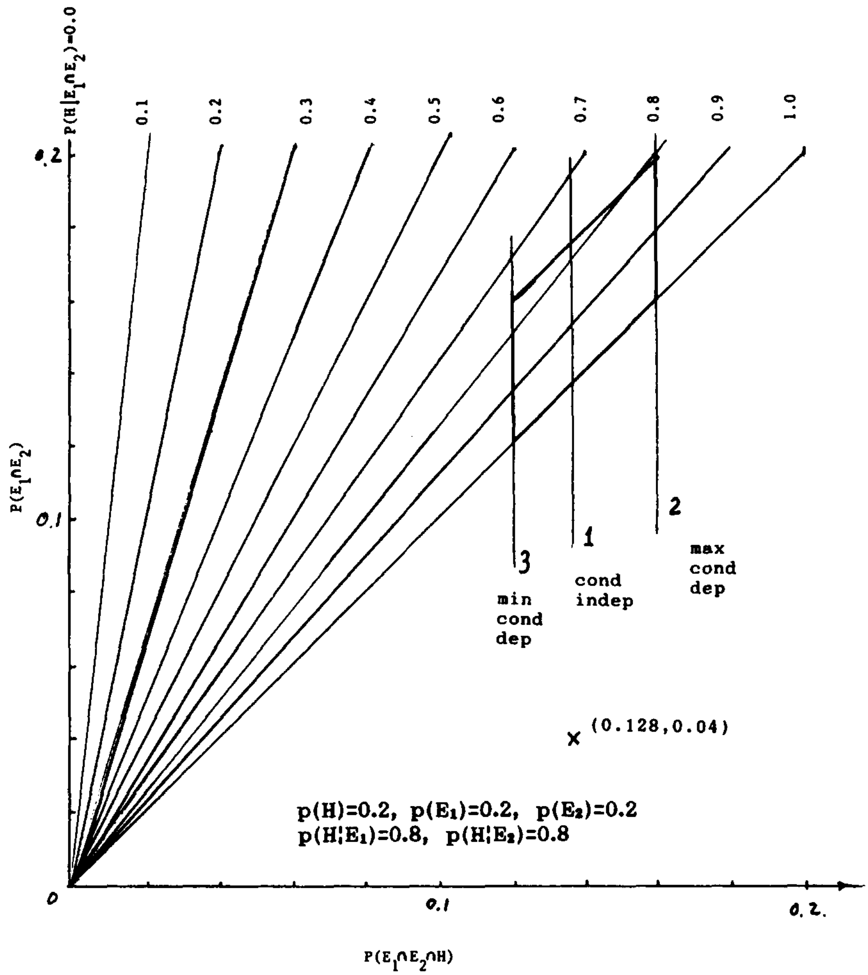


Figure 11. Effect of conditional dependency of admissible domain.

Conditional Independence

Assume

$$p(E_1 \wedge E_2 | H) = p(E_1 | H) p(E_2 | H) \tag{26}$$

The range in Eq. (16) reduces to a point value for $p(E_1 \wedge E_2 \wedge H)$, and the admissible domain degenerates to a line, which corresponds simply to the possible range of values for $p(E_1 \wedge E_2)$. This is illustrated by line 1 of Figure 11.

In this case, the consistent bounds are given by

$$\frac{\mathrm{HE12_{ind}}}{\mathrm{HE12_{ind}} + \neg\,\mathrm{HE12_{max}}} \leqslant p(H|E_1 \wedge E_2) \leqslant \frac{\mathrm{HE12_{ind}}}{\mathrm{HE12_{ind}} + \neg\,\mathrm{HE12_{min}}} \quad (27)$$

where $\mathrm{HE12_{ind}} = p(E_1|H)p(E_2|H)p(H)$.

Maximum Conditional Dependence

This condition implies that either

$$E_1 \wedge H \subseteq E_2 \wedge H \quad (28a)$$

or

$$E_1 \wedge H \supseteq E_2 \wedge H \quad (28b)$$

In this case, $p(E_1 \wedge E_2 \wedge H)$ attains its maximum possible value $\mathrm{HE12_{max}}$. The admissible domain degenerates to the right vertical line defining the domain (line 2 in Figure 11). The reduced consistent bounds are given by

$$\frac{\mathrm{HE12_{max}}}{\mathrm{HE12_{max}} + \neg\,\mathrm{HE12_{min}}} \leqslant p(H|E_1 \wedge E_2) \leqslant \frac{\mathrm{HE12_{max}}}{\mathrm{HE12_{max}} + \neg\,\mathrm{HE12_{max}}} \quad (29)$$

Minimum Conditional Dependence

Now, $p(E_1 \wedge E_2 \wedge H)$ attains its minimum possible value $\mathrm{HE12_{min}}$. The admissible domain is the left vertical line defining the domain (line 3 in Figure 11). The consistent bounds are given by

$$\frac{\mathrm{HE12_{min}}}{\mathrm{HE12_{min}} + \neg\,\mathrm{HE12_{min}}} \leqslant p(H|E_1 \wedge E_2) \leqslant \frac{\mathrm{HE12_{min}}}{\mathrm{HE12_{min}} + \neg\,\mathrm{HE12_{max}}} \quad (30)$$

Any Assumed Relation Between E_1 and E_2

The relation between E_1 and E_2 can be assumed to be

(a) Maximum dependence

$$p(E_1 \wedge E_2) = \min[p(E_1), p(E_2)] \quad (31a)$$

(b) Independence

$$p(E_1 \wedge E_2) = p(E_1)p(E_2) \quad (31b)$$

or

(c) Minimum dependence

$$p(E_1 \wedge E_2) = \max[p(E_1) + p(E_2) - 1, 0] \quad (31c)$$

However, since $p(E_1 \wedge E_2)$ interacts with $p(E_1 \wedge E_2 \wedge H)$, the constraint of Eq. (20) must be observed. In particular, since the set $E_1 \wedge E_2 \wedge H$ is contained in the set $E_1 \wedge E_2$, that is, $E_1 \wedge E_2 \wedge H \subseteq E_1 \wedge E_2$, it follows that $p(E_1 \wedge E_2 \wedge H) \leqslant p(E_1 \wedge E_2)$.

Hence, when a relation between E_1 and E_2 is assumed, $p(E_1 \wedge E_2)$ is a point value. Depending on the location of this point in the admissible domain, the admissible interval for $p(H \wedge E_1 \wedge E_2)$ can be (a) the interval $[\mathrm{HE12}_{\min}, \mathrm{HE12}_{\max}]$, as denoted by line 1 of Figure 12, (b) an interval restricted from the left, as denoted by line 2 in the same figure, or (c) an interval restricted from the right (line 3 of Figure 12). Hence, the consistent bounds can be summarized as

$$\frac{\max[\mathrm{HE12}_{\min}, p(E_1 \wedge E_2) - \neg \mathrm{HE12}_{\max}]}{p(E_1 \wedge E_2)} \leqslant p(H | E_1 \wedge E_2) \leqslant \frac{\min[\mathrm{HE12}_{\max}, p(E_1 \wedge E_2) - \neg \mathrm{HE12}_{\min}]}{p(E_1 \wedge E_2)} \tag{32a}$$

and

$$p(E_1 \wedge E_2) \geqslant \mathrm{HE12}_{\min} \tag{32b}$$

CONDITIONAL AND UNCONDITIONAL INDEPENDENCE

As mentioned in the literature review, the assumptions of unconditional independence between E_1 and E_2 and conditional independence between $E_1 \wedge H$ and $E_2 \wedge H$ are often made in contemporary evidential reasoning schemes. There are several plausible reasons for making such assumptions. First, data on the joint probability $p(E_1 \wedge E_2)$ and $p(E_1 \wedge E_2 | H)$ are usually not available, and hence some guesses must be made. Second, the independence condition corresponds to the midpoint in the spectrum of relations between two events and is considered a safe guess. Third, the independence assumptions lead to much simplified computations.

Less well known is the fact that the assumptions of unconditional independence and conditional independence may lead to inconsistent results. This can be readily shown by means of admissible domain and consistent bounds. The independence assumption corresponds to the intersection of the two lines

$$p(E_1 \wedge E_2) = p(E_1)p(E_2) \quad \text{and}$$

$$p(E_1 \wedge E_2 \wedge H) = p(E_1 | H)p(E_2 | H)p(H).$$

The point of intersection may be inside the admissible domain, or it may be outside. In the latter case, we have the situation where the set of assigned probabilities $p(H)$, . . ., $p(E_2)$ are incompatible with the independence

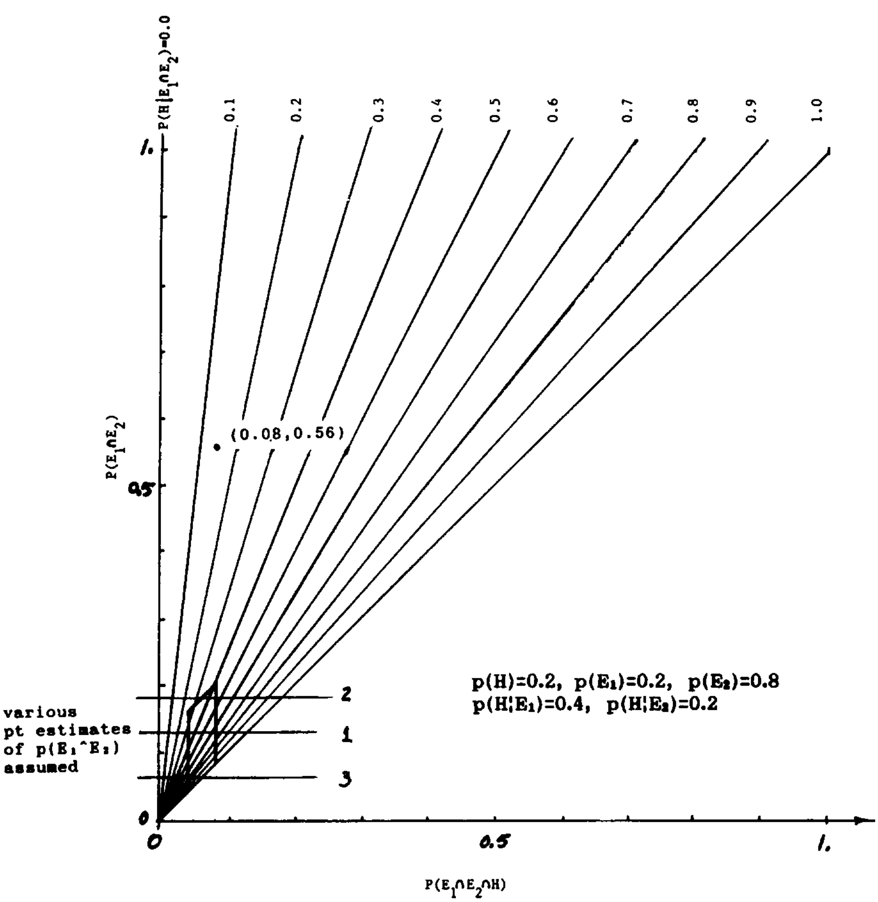


Figure 12. Effect of unconditional dependency on admissible domain.

assumptions; it is impossible to reach both independence conditions from the set of assigned probabilities.

To illustrate, suppose

$$p(H)=0.2, \qquad p(E_1)=0.2, \qquad p(E_2)=0.2,$$

$$p(H|E_1)=0.8, \qquad p(H|E_2)=0.8$$

Since

$$\begin{aligned} p(H|E_1)p(E_1) &= 0.16 \\ &< p(H) = 0.2 \\ &< p(H|E_1)p(E_1) + [1-p(E_1)] = 0.96 \end{aligned}$$

and

$$p(H|E_2)p(E_2)=0.16$$
$$<p(H)=0.2$$
$$<p(H|E_2)p(E_2)+[1-p(E_2)]=0.96$$

the set of assignments is itself consistent. Furthermore, from Bayes's theorem,

$$p(E_1|H)=\frac{p(H|E_1)p(E_1)}{p(H)}=\frac{0.8\times 0.2}{0.2}=0.8$$

and, similarly,

$$p(E_2|H)=0.8$$

If (conditional) independence between $E_1 \wedge H$ and $E_2 \wedge H$ is assumed, then

$$p(E_1 \wedge E_2 \wedge H)=p(E_1|H)p(E_2|H)p(H)$$
$$=0.8\times 0.8\times 0.2=0.128$$

Furthermore, if (unconditional) independence between E_1 and E_2 is assumed, then

$$p(E_1 \wedge E_2)=0.2\times 0.2=0.04$$

Hence, according to these independence assumptions, we have

$$p(H|E_1 \wedge E_2)=0.128/0.04=3.2$$

which is obviously incorrect. The point (0.128, 0.04) is indicated in Figure 11, and it is clear that it lies outside the admissible domain. Hence, joint use of the independence and conditional independence assumptions is not generally valid, and must be used with caution.

However, if the conditional independence assumption is made in the following way,

$$p(E_1 \wedge E_2|H)=p(E_1|H)p(E_2|H)$$

and

$$p(E_1 \wedge E_2|\neg H)=p(E_1|\neg H)p(E_2|\neg H)$$

then the fused probability will always be inside the admissible domain.

To illustrate this point, we repeat the example given previously. We have

$$p(\neg H)=0.8, \qquad p(\neg H|E_1)=0.2, \qquad p(\neg H|E_2)=0.2$$

so that

$$p(E_1|\neg H)=0.05, \qquad p(E_2|\neg H)=0.05$$

Then

$$p(H|E_1 \wedge E_2) = \frac{p(E_1 \wedge E_2 \wedge H)}{p(E_1 \wedge E_2 \wedge H) + p(E_1 \wedge E_2 \wedge \neg H)}$$

$$= \frac{0.8 \times 0.8 \times 0.2}{0.8 \times 0.8 \times 0.2 + 0.05 \times 0.05 \times 0.8}$$

$$= 0.118/0.130 = 0.98$$

which is admissible.[3]

CONSISTENT POINT ESTIMATES FOR EVIDENCE FUSION

The study of admissible domain and consistent bounds is helpful in providing insight into the nature of evidence fusion. However, working with intervals and bounds has one drawback that can be quite detrimental in applications. The problem is that the uncertainty bounds tend to increase as inference is carried from one level to another, so the final uncertainty is often so broad that it contains very little useful information.

While the broadening of uncertainty is a direct result of the accumulation of imprecise evidence, some way should be found to make the inference product easier to interpret. One way to achieve this goal is to narrow the natural bounds through explicit assumptions, for example, assumptions on dependency as discussed previously. However, we have shown the pitfall if the dependency is chosen arbitrarily.

By working with the admissible domain and choosing a point inside the domain, the corresponding dependency assumption will always be consistent with the constraints described in the previous sections. We call this point the consistent point estimate for the fused evidence to emphasize its consistency with the dependence relations. Although the consistent point estimate need only be any point inside the admissible domain, we suggest in the following three candidates that are intuitively appealing.

The first selection uses the midpoint of the bounding interval on $p(H|E_1 \wedge E_2)$. From Eqs. (23) and (24), we have

$$p(H|E_1 \wedge E_2) = (p_L + p_U)/2$$

$$= \frac{1}{2}\left\{\frac{\mathrm{HE12}_{\min}}{\mathrm{HE12}_{\min} + \neg\mathrm{HE12}_{\max}} + \frac{\mathrm{HE12}_{\max}}{\mathrm{HE12}_{\max} + \neg\mathrm{HE12}_{\min}}\right\} \tag{33}$$

We shall call this the *midpoint estimate*. The second selection uses the centroid of the admissible domain as the best estimate, that is,

[3] For fusion involving more than two mutually exclusive and exhaustive hypotheses, a discussion of which is outside the scope of this paper, conditional independence assumptions on $E_1 \wedge E_2|H_i$, $E_1 \wedge E_2|\neg H_i$, $i \geqslant 2$, will lead to the unconditional independence of E_1 and E_2 (Pednault et al. [12]).

$$p(H|E_1 \wedge E_2)=\frac{\mathrm{HE12}_{\min}+\mathrm{HE12}_{\max}}{\mathrm{HE12}_{\min}+\mathrm{HE12}_{\max}+\neg\,\mathrm{HE12}_{\min}+\neg\,\mathrm{HE12}_{\max}} \tag{34}$$

We shall call this estimate the *centroid estimate.*

The third selection is based on the assumption that all points with the admissible domain are equally good candidates. We shall call this the maximum entropy estimate.[4] Let

$$x_1=\ \mathrm{HE12}_{\min},\qquad x_2=\mathrm{HE12}_{\max}$$

$$y_1=\neg\,\mathrm{HE12}_{\min},\qquad y_2=\neg\,\mathrm{HE12}_{\max}$$

which define the opposing corners of the parallelogram domain, which has area

$$A=(x_2-x_1)(y_2-y_1)$$

Since all points inside the domain are equally likely, the uniform density distribution function is

$$f(x,y)=1/A$$

where x denotes $p(H \wedge E_1 \wedge E_2)$ and y denotes $p(E_1 \wedge E_2)$. Then the expected value of $p(H|E_1 \wedge E_2)$ is

$$p(H|E_1 \wedge E_2)=\iint \frac{x}{y} f(x,y)\,dx\,dy$$

$$=\frac{1}{A}\iint \frac{x}{y}\,dx\,dy$$

Carrying out the integration, it can be shown that when $x_1 + y_1 = 0$,

$$p(H|E_1 \wedge E_2)=\left\{\frac{x_2^2-y_2^2}{2}\ln(x_2+y_2)-\frac{x_1^2-y_2^2}{2}\ln(x_1+y_2)\right.$$
$$\left.-\frac{x_2^2-y_1^2}{2}\ln(x_2+y_1)+0.5A\right\}A^{-1} \tag{35}$$

Similarly, when $x_1 + y_1 > 0$,

$$p(H|E_1 \wedge E_2)=\left\{\frac{x_2^2-y_2^2}{2}\ln(x_2+y_2)-\frac{x_1^2-y_2^2}{2}-\ln(x_1+y_2)\right.$$
$$\left.-\frac{x_2^2-y_1^2}{2}\ln(x_2+y_1)+\frac{x_1^2-y_1^2}{2}\ln(x_1+y_1)+0.5A\right\}A^{-1} \tag{36}$$

Table 1 summarizes the point estimates computed by the midpoint, centroid, and maximum entropy methods for the cases described previously in Figures 7–12 and two additional cases. Point estimates based on the conditional and unconditional dependence assumptions are also included for comparison;

[4] Without any information on the dependency between E_1 and E_2 and between $E_1 \wedge H$ and $E_2 \wedge H$, the unbiased distribution over the domain is uniform. This is tantamount to assuming a probability distribution on the probability domains $p(E_1 \wedge E_2)$ and $p(E_1 \wedge E_2 \wedge H)$.

Table 1. Consistent Bounds and Point Estimates

							Point Estimates			
$p(H)$	$p(E_1)$	$p(E_2)$	$p(H\|E_1)$	$p(H\|E_2)$	p_L	p_U	Unconditional and Conditional Independence	Maximum Entropy	Centroid	Midpoint
0.2	0.4	0.2	0.2	0.4	0.	1.	0.4	0.42	0.4	0.5
0.4	0.8	0.4	0.4	0.8	0.75	1.	0.8	0.88	0.87	0.88
0.6	0.8	0.4	0.6	0.2	0.	0.25	0.2	0.12	0.13	0.13
0.8	0.6	0.8	0.8	0.8	0.73	0.86	0.8	0.80	0.80	0.80
0.2	0.2	0.2	0.8	0.8	0.75	1.	~~3.2~~	0.88	0.87	0.88
0.2	0.2	0.8	0.4	0.2	0.25	1.	0.4	0.55	0.5	0.63
0.2	0.2	0.6	0.2	0.2	0.	1.	0.2	0.26	0.2	0.50
0.4	0.4	0.4	0.6	0.8	0.67	1.	~~1.6~~	0.84	0.83	0.84

violations are indicated as crossed-out numbers. As can be seen from these numerical results, which are given to two significant figures, all three point estimates produce similar results that are, by the nature of the domain from which they are derived, guaranteed to be admissible. Furthermore, the centroid and maximum entropy estimates are very close to each other. Hence, the centroid estimate may be more appealing due to its computational simplicity.

We emphasize that the use of intervals for probabilistic evidence, while perhaps not yielding much information in some circumstances because of the nature of the problem, is correct and does not make any unwarranted assumptions. If the bounds are very wide, then they only reflect certain inherent uncertainty about the problem situation. All one can really do is try to obtain more data. The use of point estimates discussed in this section is not intended to ignore the need for more information. Point estimates are often discussed in the literature, and they are included here for completeness.

CONCLUSION

The evidence fusion process has been investigated from the point of view of probability theory. The study has focused on a crucial but seldom addressed issue: that the uncertainty bounds for the constituents of the Bayesian updating equation are interactive. Consequently, although methodologies on the computation of the uncertainty bounds of the constituents are available, notably Nilsson's probabilistic logic world, computation of the fused evidence is still very difficult.

Using a simple system consisting of a single hypothesis and two evidence sources, we have delineated the nature of the interaction and how it affects the estimate of the fused evidence. This is possible because the bounds on the constituents and their interaction can be derived analytically. Different assumptions of dependency between the constituent evidence will, in general, lead to narrower bounds. However, the validity of the assumptions and their consistency with the assigned probabilities should always be checked. It is shown, by means of the admissible domain, how commonly used conditional independence assumptions may lead to inconsistent results.

For the more general and complex cases, an analytical derivation appears very difficult. While the bounds on the constituents can be computed by brute-force numerical techniques such as linear programming in conjunction with Nilsson's possible world, the Bayes function is nonlinear and hence not amenable to the linear programming technique. This task is the ultimate objective of our study, and the development described in this paper is a first step toward that goal.

ACKNOWLEDGMENT

We thank all the referees who took time to comment frankly on the paper, provide encouragement, and more important, offer very constructive criticism.

We believe that because of their effort the quality of the paper and our knowledge of the subject have been greatly improved.

References

1. Shortliffe, E., *Computer-Based Medical Consultations: MYCIN,* American Elsevier, New York, 1976.
2. Duda, R., Hart, P., and Nilsson, N., Subjective Bayesian methods for rule-based inference systems, Proc. National Computer Conference, New York, 1976.
3. Dempster, A., Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat.,* 38, 325–429, 1967.
4. Shafer, G., *A Mathematical Theory of Evidence,* Princeton Univ. Press, Princeton, N.J., 1976.
5. Heckerman, D., Probabilities interpretations for MYCIN's certainty factors, in *Uncertainty in Artificial Intelligence* (L. N. Kanal and J. F. Lemmer, Eds.), Elsevier, New York, 167–196, 1986.
6. Zadeh, L. A., Review of Shafer's *A Mathematical Theory of Evidence, AI Mag.* 5(3), 81–83, 1984.
7. Zadeh, L. A., A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, *AI Mag.* 7(2), 85–90, 1986.
8. Hunter, D., Dempster-Shafer vs probabilistic logic, Proc. Third Workshop on Uncertainty in Artificial Intelligence, AAAI, Menlo Park, 22–29, 1987.
9. Yen, J., Can evidence be combined in the Dempster-Shafer theory, Proc. Third Workshop on Uncertainty in Artificial Intelligence, AAAI, Menlo Park, 70–76, 1987.
10. Good, I. J., Subjective probability as the measure of a non-measurable set, in *Good Thinking,* Univ. Minnesota Press, Minneapolis, 1983.
11. Cooper, G., A diagnostic method that uses causal knowledge and linear programming in the application of Bayes' formula, *Computer Methods Programs Biomed.* 22, 1986.
12. Pednault, E., Zucker, S., and Muresan, L., On the independence assumption underlying subjective Bayesian updating, *AI* 16, 213–222, 1981.
13. Glymour, C., Independence assumptions and Bayesian updating, *AI* 25, 95–99, 1985.
14. Johnson, R. W., Independence and Bayesian updating, in *Uncertainty in Artificial Intelligence* (L. N. Kanal and J. F. Lemmer, Eds.), Elsevier, New York, 197–201, 1986.
15. Nilsson, N. J. Probabilistic logic, *AI,* 28, 71–87, 1986.